Matrix Computation

To set up a philosophy against physics is rash;
philosophers who have done so
have always ended in disaster.
— Bertrand Russell

Definitions and Basic Results

- Let $A \equiv [\, a_{ij} \,]_{1 \leq i \leq m, 1 \leq j \leq n}$, or simply $A \in \boldsymbol{R}^{m \times n}$, denote an $m \times n$ matrix.

- It can also be represented as $[\, a_1, a_2, \dots, a_n \,]$ where $a_i \in \boldsymbol{R}^m$ are vectors.

  – Vectors are column vectors unless stated otherwise.

- A is a square matrix when $m = n$.

- The rank of a matrix is the largest number of linearly independent columns.

Definitions and Basic Results (continued)

- A square matrix A is said to be symmetric if $A^{\mathrm{T}} = A$.

- A real $n \times n$ matrix

$$A \equiv [\, a_{ij} \,]_{i,j}$$

  is diagonally dominant if $|\, a_{ii} \,| > \sum_{j \neq i} |\, a_{ij} \,|$ for $1 \leq i \leq n$.

  - Such matrices are nonsingular.

- The identity matrix is the square matrix

$$I \equiv \mathrm{diag}[\, 1, 1, \ldots, 1 \,].$$

Definitions and Basic Results (concluded)

- A matrix has full column rank if its columns are linearly independent.

- A real symmetric matrix A is positive definite if

$$x^{\mathrm{T}}Ax = \sum_{i,j} a_{ij}x_i x_j > 0$$

  for any nonzero vector x.

- A matrix A is positive definite if and only if there exists a matrix W such that $A = W^{\mathrm{T}}W$ and W has full column rank.

Cholesky Decomposition

- Positive definite matrices can be factored as

$$A = LL^{\mathrm{T}},$$

  called the Cholesky decomposition.

  − Above, L is a lower triangular matrix.

Generation of Multivariate Distribution

- Let $\boldsymbol{x} \equiv [\, x_1, x_2, \ldots, x_n \,]^{\mathrm{T}}$ be a vector random variable with a positive definite covariance matrix C.

- As usual, assume $E[\, \boldsymbol{x} \,] = \boldsymbol{0}$.

- This covariance structure can be matched by $P\boldsymbol{y}$.

  - $C = PP^{\mathrm{T}}$ is the Cholesky decomposition of C.[a]

  - $\boldsymbol{y} \equiv [\, y_1, y_2, \ldots, y_n \,]^{\mathrm{T}}$ is a vector random variable with a covariance matrix equal to the identity matrix.

---

[a]What if C is not positive definite? See Lai (`R93942114`) and Lyuu (2007).

Generation of Multivariate Normal Distribution

- Suppose we want to generate the multivariate normal distribution with a covariance matrix $C = PP^{\mathrm{T}}$.

- We start with independent standard normal distributions $y_1, y_2, \ldots, y_n$.

- Then $P[\, y_1, y_2, \ldots, y_n \,]^{\mathrm{T}}$ has the desired distribution.

Multivariate Derivatives Pricing

- Generating the multivariate normal distribution is essential for the Monte Carlo pricing of multivariate derivatives (p. 629).

- For example, the rainbow option on k assets has payoff

$$\max(\max(S_1, S_2, \ldots, S_k) - X, 0)$$

  at maturity.

- The closed-form formula is a multi-dimensional integral.[a]

---

[a]Johnson (1987); Chen (`D95723006`) and Lyuu (2009).

Multivariate Derivatives Pricing (concluded)

- Suppose $dS_j / S_j = r\,dt + \sigma_j\,dW_j$, $1 \le j \le k$, where C is the correlation matrix for $dW_1, dW_2, \ldots, dW_k$.

- Let $C = PP^{\mathrm{T}}$.

- Let ξ consist of k independent random variables from $N(0, 1)$.

- Let $\xi' = P\xi$.

- Similar to Eq. (78) on p. 669,

$$S_{i+1} = S_i e^{(r - \sigma_j^2/2)\,\Delta t + \sigma_j \sqrt{\Delta t}\,\xi_j'}, \quad 1 \le j \le k.$$

Least-Squares Problems

- The least-squares (LS) problem is concerned with

$$\min_{x \in R^n} \parallel Ax - b \parallel,$$

  where $A \in \boldsymbol{R}^{m \times n}$, $b \in \boldsymbol{R}^m$, $m \geq n$.

- The LS problem is called regression analysis in statistics and is equivalent to minimizing the mean-square error.

- Often written as

$$Ax = b.$$

Polynomial Regression

- In polynomial regression, $x_0 + x_1 x + \cdots + x_n x^n$ is used to fit the data $\{(a_1, b_1), (a_2, b_2), \ldots, (a_m, b_m)\}$.

- This leads to the LS problem,

$$
\begin{bmatrix}
1 & a_1 & a_1^2 & \cdots & a_1^n \\
1 & a_2 & a_2^2 & \cdots & a_2^n \\
\vdots & \vdots & \vdots & \ddots & \vdots \\
1 & a_m & a_m^2 & \cdots & a_m^n
\end{bmatrix}
\begin{bmatrix}
x_0 \\
x_1 \\
\vdots \\
x_n
\end{bmatrix}
=
\begin{bmatrix}
b_1 \\
b_2 \\
\vdots \\
b_m
\end{bmatrix}.
$$

- Consult the text for solutions.

American Option Pricing by Simulation

- The continuation value of an American option is the conditional expectation of the payoff from keeping the option alive now.

- The option holder must compare the immediate exercise value and the continuation value.

- In standard Monte Carlo simulation, each path is treated independently of other paths.

- But the decision to exercise the option cannot be reached by looking at only one path alone.

The Least-Squares Monte Carlo Approach

- The continuation value can be estimated from the cross-sectional information in the simulation by using least squares.[a]

- The result is a function (of the state) for estimating the continuation values.

- Use the function to estimate the continuation value for each path to determine its cash flow.

- This is called the least-squares Monte Carlo (LSM) approach and is provably convergent.[b]

---

[a]Longstaff and Schwartz (2001).
[b]Clément, Lamberton, and Protter (2002).

A Numerical Example

- Consider a 3-year American put on a non-dividend-paying stock.

- The put is exercisable at years 0, 1, 2, and 3.

- The strike price $X = 105$.

- The annualized riskless rate is $r = 5\%$.

- The spot stock price is 101.

  – The annual discount factor hence equals 0.951229.

- We use only 8 price paths to illustrate the algorithm.

A Numerical Example (continued)

Stock price paths

Path	Year 0	Year 1	Year 2	Year 3
1	**101**	**97.6424**	**92.5815**	107.5178
2	**101**	**101.2103**	105.1763	**102.4524**
3	**101**	105.7802	**103.6010**	124.5115
4	**101**	**96.4411**	**98.7120**	108.3600
5	**101**	124.2345	**101.0564**	**104.5315**
6	**101**	**95.8375**	**93.7270**	**99.3788**
7	**101**	108.9554	**102.4177**	**100.9225**
8	**101**	**104.1475**	113.2516	115.0994
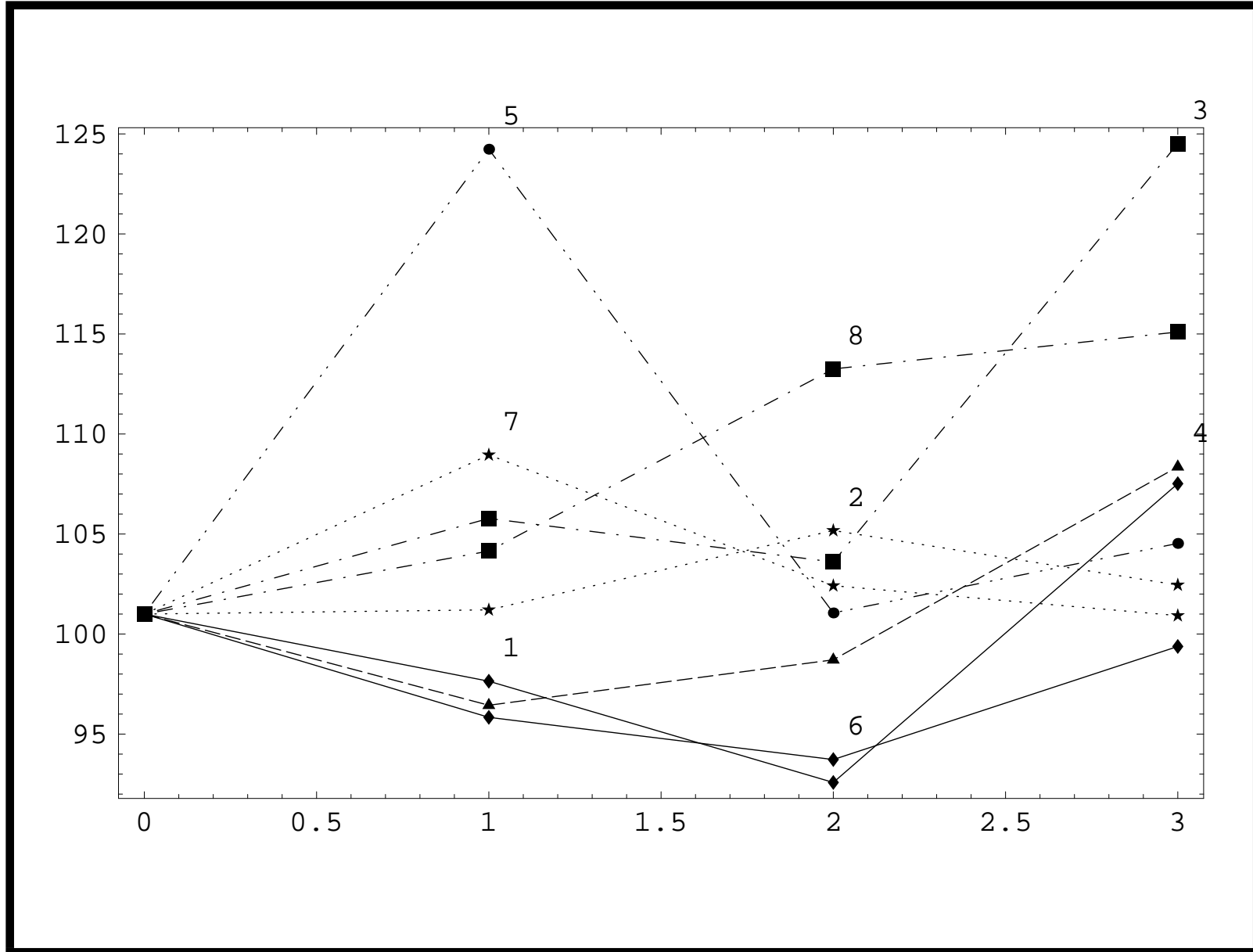
A Numerical Example (continued)

- We use the basis functions $1, x, x^2$.

    - Other basis functions are possible.[a]

- The plot next page shows the final estimated optimal exercise strategy given by LSM.

- We now proceed to tackle our problem.

- The idea is to calculate the cash flow along each path, using information from all paths.

---

[a]Laguerre polynomials, Hermite polynomials, Legendre polynomials, Chebyshev polynomials, Gedenbauer polynomials, and Jacobi polynomials.
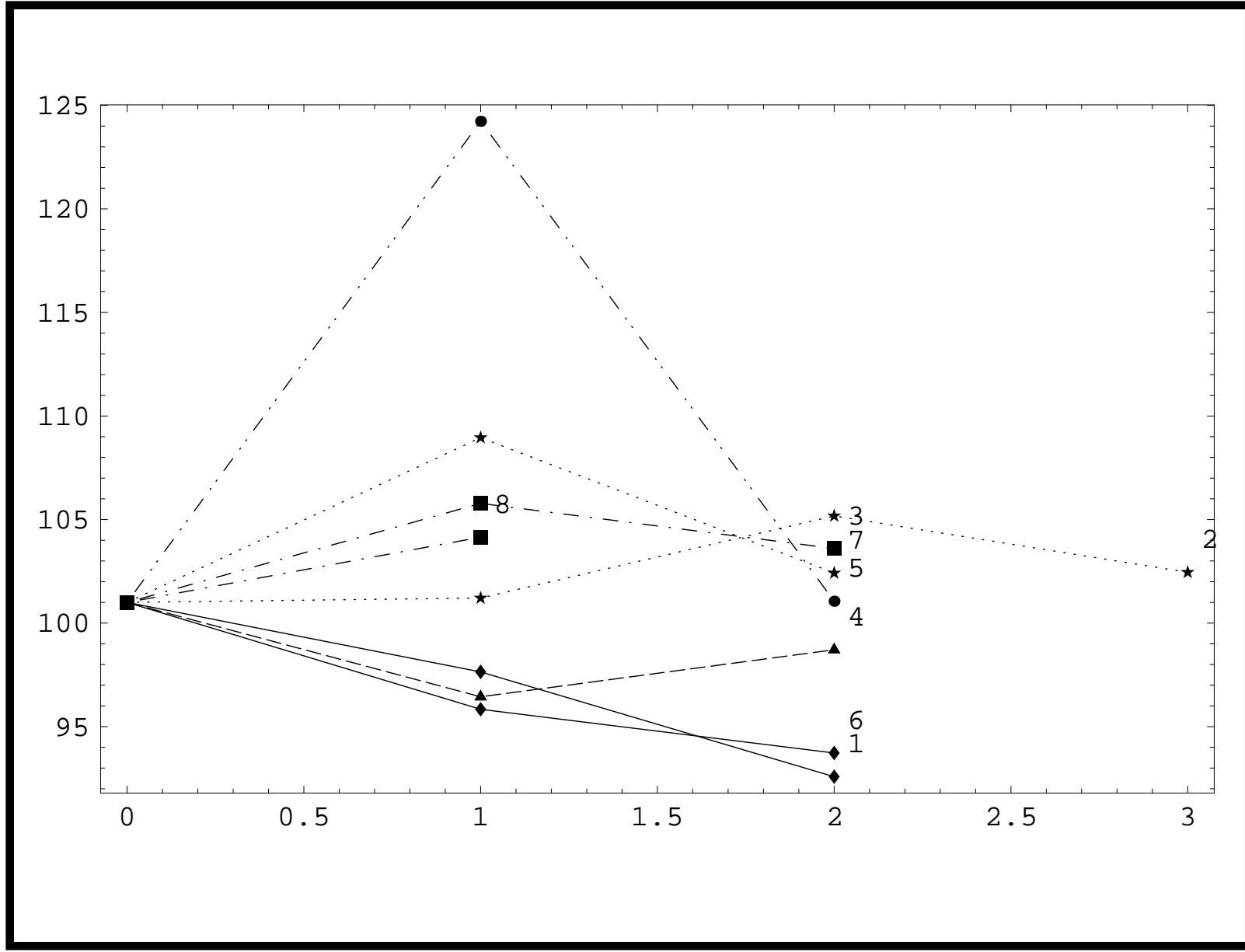
A Numerical Example (continued)

Cash flows at year 3

Path	Year 0	Year 1	Year 2	Year 3
1	—	—	—	0
2	—	—	—	2.5476
3	—	—	—	0
4	—	—	—	0
5	—	—	—	0.4685
6	—	—	—	5.6212
7	—	—	—	4.0775
8	—	—	—	0

A Numerical Example (continued)

- The cash flows at year 3 are the exercise value if the put is in the money.

- Only 4 paths are in the money: 2, 5, 6, 7.

- Some of the cash flows may not occur if the put is exercised earlier, which we will find out step by step.

- Incidentally, the European counterpart has a value of

$$0.951229^3 \times \frac{2.5476 + 0.4685 + 5.6212 + 4.0775}{8} = 1.3680.$$

A Numerical Example (continued)

- We move on to year 2.

- For each state that is in the money at year 2, we must decide whether to exercise it.

- There are 6 paths for which the put is in the money: 1, 3, 4, 5, 6, 7.

- Only in-the-money paths will be used in the regression because they are where early exercise is relevant.

  - If there were none, we would move on to year 1.

A Numerical Example (continued)

- Let x denote the stock prices at year 2 for those 6 paths.

- Let y denote the corresponding discounted future cash flows (at year 3) if the put is not exercised at year 2.

A Numerical Example (continued)

Regression at year 2

Path	x	y
1	92.5815	0×0.951229
2	—	—
3	103.6010	0×0.951229
4	98.7120	0×0.951229
5	101.0564	0.4685×0.951229
6	93.7270	5.6212×0.951229
7	102.4177	4.0775×0.951229
8	—	—

A Numerical Example (continued)

- We regress y on 1, x, and x^2.

- The result is

$$f(x) = 22.08 - 0.313114 \times x + 0.00106918 \times x^2.$$

- f estimates the continuation value conditional on the stock price at year 2.

- We next compare the immediate exercise value and the continuation value.

A Numerical Example (continued)

Optimal early exercise decision at year 2

Path	Exercise	Continuation
1	12.4185	$f(92.5815) = 2.2558$
2	—	—
3	1.3990	$f(103.6010) = 1.1168$
4	6.2880	$f(98.7120) = 1.5901$
5	3.9436	$f(101.0564) = 1.3568$
6	11.2730	$f(93.7270) = 2.1253$
7	2.5823	$f(102.4177) = 0.3326$
8	—	—

A Numerical Example (continued)

- Amazingly, the put should be exercised in all 6 paths: 1, 3, 4, 5, 6, 7.

- Now, any positive cash flow at year 3 should be set to zero for these paths as the put is exercised before year 3.

  – They are paths 5, 6, 7.

- Hence the cash flows on p. 726 become the next ones.

A Numerical Example (continued)

Cash flows at years 2 & 3

Path	Year 0	Year 1	Year 2	Year 3
1	—	—	12.4185	0
2	—	—	0	2.5476
3	—	—	1.3990	0
4	—	—	6.2880	0
5	—	—	3.9436	0
6	—	—	11.2730	0
7	—	—	2.5823	0
8	—	—	0	0

A Numerical Example (continued)

- We move on to year 1.

- For each state that is in the money at year 1, we must decide whether to exercise it.

- There are 5 paths for which the put is in the money: 1, 2, 4, 6, 8.

- Only in-the-money paths will be used in the regression because they are where early exercise is relevant.

  - If there were none, we would move on to year 0.

A Numerical Example (continued)

- Let x denote the stock prices at year 1 for those 5 paths.

- Let y denote the corresponding discounted future cash flows if the put is not exercised at year 1.

- From p. 734, we have the following table.

A Numerical Example (continued)

Regression at year 1

Path	x	y
1	97.6424	12.4185×0.951229
2	101.2103	2.5476×0.951229^2
3	—	—
4	96.4411	6.2880×0.951229
5	—	—
6	95.8375	11.2730×0.951229
7	—	—
8	104.1475	0

A Numerical Example (continued)

- We regress y on 1, x, and x^2.

- The result is

$$f(x) = -420.964 + 9.78113 \times x - 0.0551567 \times x^2.$$

- f estimates the continuation value conditional on the stock price at year 1.

- We next compare the immediate exercise value and the continuation value.

A Numerical Example (continued)

Optimal early exercise decision at year 1

Path	Exercise	Continuation
1	7.3576	$f(97.6424) = 8.2230$
2	3.7897	$f(101.2103) = 3.9882$
3	—	—
4	8.5589	$f(96.4411) = 9.3329$
5	—	—
6	9.1625	$f(95.8375) = 9.83042$
7	—	—
8	0.8525	$f(104.1475) = -0.551885$

A Numerical Example (continued)

- The put should be exercised for 1 path only: 8.

- Now, any positive future cash flow should be set to zero for this path as the put is exercised before years 2 and 3.

  – But there is none.

- Hence the cash flows on p. 734 become the next ones.

- They also confirm the plot on p. 725.

A Numerical Example (continued)

Cash flows at years 1, 2, & 3

Path	Year 0	Year 1	Year 2	Year 3
1	—	0	12.4185	0
2	—	0	0	2.5476
3	—	0	1.3990	0
4	—	0	6.2880	0
5	—	0	3.9436	0
6	—	0	11.2730	0
7	—	0	2.5823	0
8	—	0.8525	0	0

A Numerical Example (continued)

- We move on to year 0.

- The continuation value is, from p 741,

$$(12.4185 \times 0.951229^2 + 2.5476 \times 0.951229^3$$
$$+1.3990 \times 0.951229^2 + 6.2880 \times 0.951229^2$$
$$+3.9436 \times 0.951229^2 + 11.2730 \times 0.951229^2$$
$$+2.5823 \times 0.951229^2 + 0.8525 \times 0.951229)/8$$
$$= 4.66263.$$

A Numerical Example (concluded)

- As this is larger than the immediate exercise value of $105 - 101 = 4$, the put should not be exercised at year 0.

- Hence the put's value is estimated to be 4.66263.

- Compare this to the European put's value of 1.3680 (p. 727).

- Why is the LSM estimate a lower bound?[a]

---

[a]Contributed by Mr. Yang, Jui-Chung (`D97723002`) on April 29, 2009.

Time Series Analysis

The historian is a prophet in reverse.

— Friedrich von Schlegel (1772–1829)

Conditional Variance Models for Price Volatility

- Although a stationary model (see text for definition) has constant variance, its *conditional* variance may vary.

- Take for example an AR(1) process $X_t = aX_{t-1} + \epsilon_t$ with $|a| < 1$.

  - Here, ϵ_t is a stationary, uncorrelated process with zero mean and constant variance σ^2.

- The conditional variance,

$$\text{Var}[\, X_t \mid X_{t-1}, X_{t-2}, \ldots\,],$$

equals σ^2, which is smaller than the *unconditional* variance $\text{Var}[\, X_t\,] = \sigma^2/(1 - a^2)$.

Conditional Variance Models for Price Volatility (concluded)

- In the lognormal model, the conditional variance evolves independently of past returns.

- Suppose we assume that conditional variances are deterministic functions of past returns:

$$V_t = f(X_{t-1}, X_{t-2}, \dots)$$

  for some function f.

- Then V_t can be computed given the information set of past returns:

$$I_{t-1} \equiv \{\, X_{t-1}, X_{t-2}, \dots \}.$$

ARCH Models[a]

- An influential model in this direction is the
  autoregressive conditional heteroskedastic (ARCH)
  model.

- Assume that $\{ U_t \}$ is a Gaussian stationary,
  uncorrelated process.

---

[a]Engle (1982), co-winner of the 2003 Nobel Prize in Economic Sciences.

ARCH Models (continued)

- The ARCH(p) process is defined by

$$X_t - \mu = \left(a_0 + \sum_{i=1}^{p} a_i (X_{t-i} - \mu)^2 \right)^{1/2} U_t,$$

  where $a_1, \ldots, a_p \geq 0$ and $a_0 > 0$.
  - Thus $X_t \,|\, I_{t-1} \sim N(\mu, V_t^2)$.

- The variance V_t^2 satisfies

$$V_t^2 = a_0 + \sum_{i=1}^{p} a_i (X_{t-i} - \mu)^2.$$

- The volatility at time t as estimated at time $t - 1$ depends on the p most recent observations on squared returns.

ARCH Models (concluded)

- The ARCH(1) process

$$X_t - \mu = (a_0 + a_1(X_{t-1} - \mu)^2)^{1/2} U_t$$

  is the simplest.

- For it,

$$\mathrm{Var}[\, X_t \mid X_{t-1} = x_{t-1} \,] = a_0 + a_1(x_{t-1} - \mu)^2.$$

- The process $\{\, X_t \,\}$ is stationary with finite variance if and only if $a_1 < 1$, in which case $\mathrm{Var}[\, X_t \,] = a_0/(1 - a_1)$.

GARCH Models[a]

- A very popular extension of the ARCH model is the generalized autoregressive conditional heteroskedastic (GARCH) process.

- The simplest GARCH$(1, 1)$ process adds $a_2 V_{t-1}^2$ to the ARCH(1) process, resulting in

$$V_t^2 = a_0 + a_1 (X_{t-1} - \mu)^2 + a_2 V_{t-1}^2.$$

- The volatility at time t as estimated at time $t - 1$ depends on the squared return and the estimated volatility at time $t - 1$.

---

[a]Bollerslev (1986); Taylor (1986).

GARCH Models (concluded)

- The estimate of volatility averages past squared returns by giving heavier weights to recent squared returns (see text).

- It is usually assumed that $a_1 + a_2 < 1$ and $a_0 > 0$, in which case the unconditional, long-run variance is given by $a_0/(1 - a_1 - a_2)$.

- A popular special case of $\mathrm{GARCH}(1, 1)$ is the exponentially weighted moving average process, which sets a_0 to zero and a_2 to $1 - a_1$.

- This model is used in J.P. Morgan's RiskMetrics™.

GARCH Option Pricing[a]

- Options can be priced when the underlying asset's return follows a GARCH process.

- Let S_t denote the asset price at date t.

- Let h_t^2 be the conditional variance of the return over the period $[\, t, t+1 \,]$ given the information at date t.

  - "One day" is merely a convenient term for any elapsed time Δt.

---

[a]A Bloomberg quant said to me on Feb 29, 2008, that GARCH option pricing is seldom used in trading.

GARCH Option Pricing (continued)

- Adopt the following risk-neutral process for the price dynamics:[a]

$$\ln \frac{S_{t+1}}{S_t} = r - \frac{h_t^2}{2} + h_t \epsilon_{t+1}, \tag{81}$$

where

$$
\begin{aligned}
h_{t+1}^2 &= \beta_0 + \beta_1 h_t^2 + \beta_2 h_t^2 (\epsilon_{t+1} - c)^2, \tag{82} \\
\epsilon_{t+1} &\sim N(0, 1) \text{ given information at date } t, \\
r &= \text{daily riskless return,} \\
c &\geq 0.
\end{aligned}
$$

---

[a]Duan (1995).

GARCH Option Pricing (continued)

- The five unknown parameters of the model are c, h_0, β_0, β_1, and β_2.

- It is postulated that $\beta_0, \beta_1, \beta_2 \geq 0$ to make the conditional variance positive.

- The above process, called the nonlinear asymmetric GARCH model, generalizes the $\text{GARCH}(1,1)$ model (see text).

GARCH Option Pricing (continued)

- It captures the volatility clustering in asset returns first noted by Mandelbrot (1963).[a]

  - When $c = 0$, a large ϵ_{t+1} results in a large h_{t+1}, which in turns tends to yield a large h_{t+2}, and so on.

- It also captures the negative correlation between the asset return and changes in its (conditional) volatility.[b]

  - For $c > 0$, a positive ϵ_{t+1} (good news) tends to decrease h_{t+1}, whereas a negative ϵ_{t+1} (bad news) tends to do the opposite.

---

[a] "... large changes tend to be followed by large changes—of either sign—and small changes tend to be followed by small changes ..."

[b] Noted by Black (1976): Volatility tends to rise in response to "bad news" and fall in response to "good news."

GARCH Option Pricing (concluded)

- With $y_t \equiv \ln S_t$ denoting the logarithmic price, the model becomes

$$y_{t+1} = y_t + r - \frac{h_t^2}{2} + h_t \epsilon_{t+1}. \tag{83}$$

- The pair (y_t, h_t^2) completely describes the current state.

- The conditional mean and variance of y_{t+1} are clearly

$$
\begin{aligned}
E[\, y_{t+1} \,|\, y_t, h_t^2 \,] &= y_t + r - \frac{h_t^2}{2}, &\tag{84} \\
\mathrm{Var}[\, y_{t+1} \,|\, y_t, h_t^2 \,] &= h_t^2. &\tag{85}
\end{aligned}
$$

GARCH Model: Inferences

- Suppose the parameters c, h_0, β_0, β_1, and β_2 are given.

- Then we can recover $h_1, h_2, \ldots, h_n$ and $\epsilon_1, \epsilon_2, \ldots, \epsilon_n$ from the prices

$$S_0, S_1, \ldots, S_n$$

  under the GARCH model (81) on p. 754.

- This property is useful in statistical inferences.

The Ritchken-Trevor (RT) Algorithm[a]

- The GARCH model is a continuous-state model.

- To approximate it, we turn to trees with *discrete* states.

- Path dependence in GARCH makes the tree for asset prices explode exponentially (why?).

- We need to mitigate this combinatorial explosion.

---

[a]Ritchken and Trevor (1999).

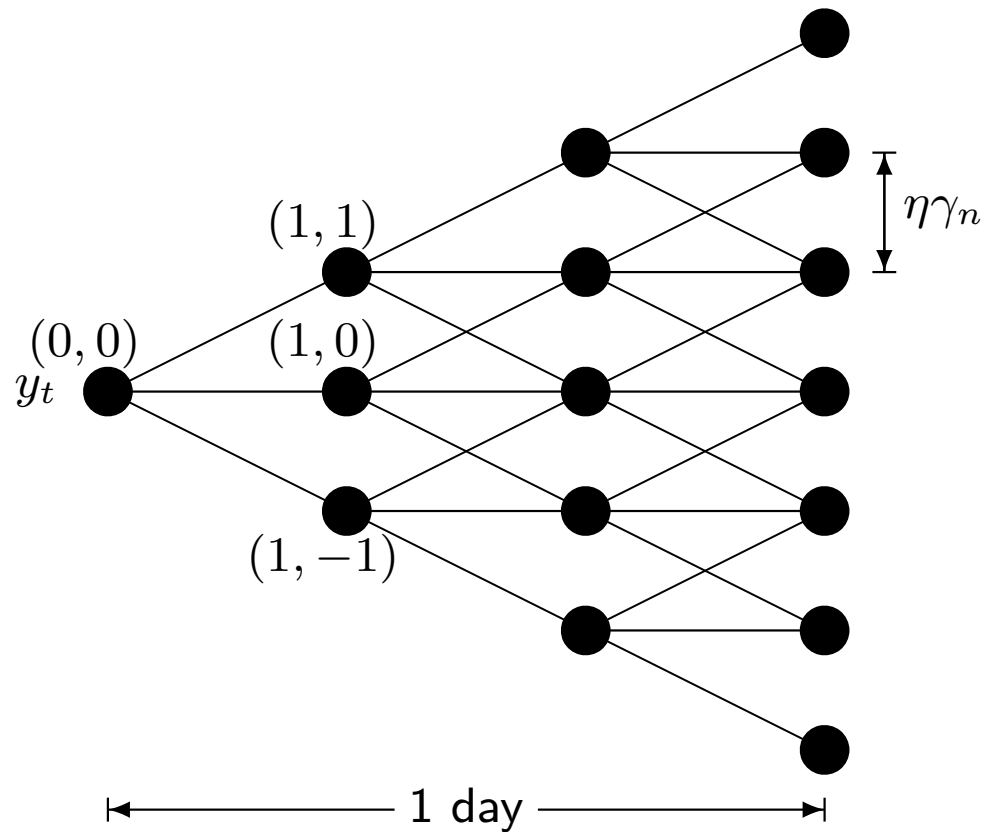# The Ritchken-Trevor Algorithm (continued)

- Partition a day into n periods.

- Three states follow each state (y_t, h_t^2) after a period.

- As the trinomial model combines, $2n + 1$ states at date $t + 1$ follow each state at date t (recall p. 584).

- These $2n + 1$ values must approximate the distribution of (y_{t+1}, h_{t+1}^2).

- So the conditional moments (84)–(85) at date $t + 1$ on p. 757 must be matched by the trinomial model to guarantee convergence to the continuous-state model.

The Ritchken-Trevor Algorithm (continued)

- It remains to pick the jump size and the three branching probabilities.

- The role of σ in the Black-Scholes option pricing model is played by h_t in the GARCH model.

- As a jump size proportional to $\sigma/\sqrt{n}$ is picked in the BOPM, a comparable magnitude will be chosen here.

- Define $\gamma \equiv h_0$, though other multiples of h_0 are possible, and

$$\gamma_n \equiv \frac{\gamma}{\sqrt{n}}.$$

- The jump size will be some integer multiple η of γ_n.

- We call η the jump parameter (p. 762).

The seven values on the right approximate the distribution of logarithmic price y_{t+1}.

The Ritchken-Trevor Algorithm (continued)

- The middle branch does not change the underlying asset's price.

- The probabilities for the up, middle, and down branches are

$$
\begin{aligned}
p_u &= \frac{h_t^2}{2\eta^2\gamma^2} + \frac{r - (h_t^2/2)}{2\eta\gamma\sqrt{n}}, & (86) \\
p_m &= 1 - \frac{h_t^2}{\eta^2\gamma^2}, & (87) \\
p_d &= \frac{h_t^2}{2\eta^2\gamma^2} - \frac{r - (h_t^2/2)}{2\eta\gamma\sqrt{n}}. & (88)
\end{aligned}
$$

The Ritchken-Trevor Algorithm (continued)

- It can be shown that:

  - The trinomial model takes on $2n + 1$ values at date $t + 1$ for y_{t+1}.

  - These values have a matching mean for y_{t+1}.

  - These values have an asymptotically matching variance for y_{t+1}.

- The central limit theorem guarantees the desired convergence as n increases.

The Ritchken-Trevor Algorithm (continued)

- We can dispense with the intermediate nodes *between* dates to create a $(2n + 1)$-nomial tree (p. 766).

- The resulting model is multinomial with $2n + 1$ branches from any state (y_t, h_t^2).

- There are two reasons behind this manipulation.
  - Interdate nodes are created merely to approximate the continuous-state model after one day.
  - Keeping the interdate nodes results in a tree that can be as much as n times larger.[a]

  _____

  [a]Contrast that with the case on p. 330.

y_t

$\eta\gamma_n$

|← ——————— 1 day ——————— →|
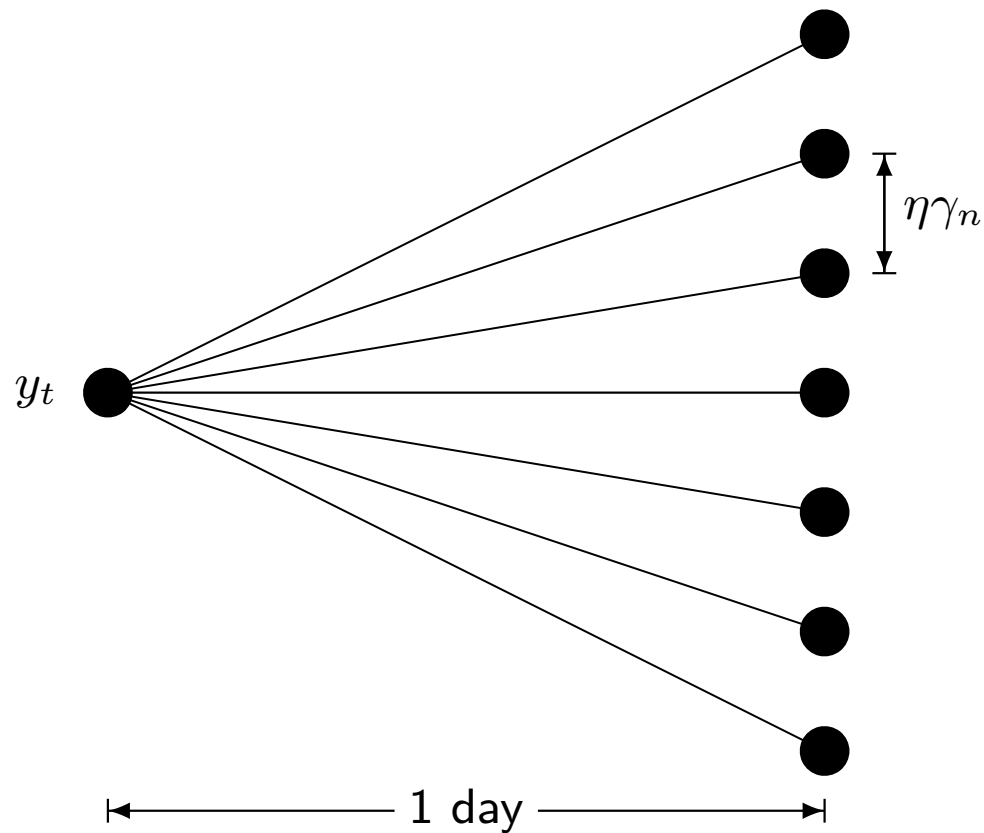
This heptanomial tree is the outcome of the trinomial tree on p. 762 after its intermediate nodes are removed.

The Ritchken-Trevor Algorithm (continued)

- A node with logarithmic price $y_t + \ell \eta \gamma_n$ at date $t+1$ follows the current node at date t with price y_t, where $-n \leq \ell \leq n$.

- To reach that price in n periods, the number of up moves must exceed that of down moves by exactly ℓ.

- The probability that this happens is

$$P(\ell) \equiv \sum_{j_u, j_m, j_d} \frac{n!}{j_u! \, j_m! \, j_d!} \, p_u^{j_u} p_m^{j_m} p_d^{j_d},$$

with $j_u, j_m, j_d \geq 0$, $n = j_u + j_m + j_d$, and $\ell = j_u - j_d$.

The Ritchken-Trevor Algorithm (continued)

- A particularly simple way to calculate the $P(\ell)$s starts by noting that

$$(p_u x + p_m + p_d x^{-1})^n = \sum_{\ell=-n}^{n} P(\ell)\, x^{\ell}. \qquad (89)$$

 – Convince yourself that this trick does the "accounting" correctly.

- So we expand $(p_u x + p_m + p_d x^{-1})^n$ and retrieve the probabilities by reading off the coefficients.

- It can be computed in $O(n^2)$ time.

The Ritchken-Trevor Algorithm (continued)

- The updating rule (82) on p. 754 must be modified to account for the adoption of the discrete-state model.

- The logarithmic price $y_t + \ell\eta\gamma_n$ at date $t+1$ following state (y_t, h_t^2) at date t has a variance equal to

$$h_{t+1}^2 = \beta_0 + \beta_1 h_t^2 + \beta_2 h_t^2 (\epsilon'_{t+1} - c)^2, \qquad (90)$$

  - Above,

$$\epsilon'_{t+1} = \frac{\ell\eta\gamma_n - (r - h_t^2/2)}{h_t}, \quad \ell = 0, \pm 1, \pm 2, \ldots, \pm n,$$

    is a discrete random variable with $2n+1$ values.

The Ritchken-Trevor Algorithm (continued)

- Different conditional variances h_t^2 may require different η so that the probabilities calculated by Eqs. (86)–(88) on p. 763 lie between 0 and 1.

- This implies varying jump sizes.

- The necessary requirement $p_m \geq 0$ implies $\eta \geq h_t/\gamma$.

- Hence we try

$$\eta = \lceil h_t/\gamma \rceil, \lceil h_t/\gamma \rceil + 1, \lceil h_t/\gamma \rceil + 2, \ldots$$

 until valid probabilities are obtained or until their nonexistence is confirmed.

The Ritchken-Trevor Algorithm (continued)

- The sufficient and necessary condition for valid probabilities to exist is[a]

$$\frac{|\, r - (h_t^2/2)\,|}{2\eta\gamma\sqrt{n}} \leq \frac{h_t^2}{2\eta^2\gamma^2} \leq \min\left(1 - \frac{|\, r - (h_t^2/2)\,|}{2\eta\gamma\sqrt{n}}, \frac{1}{2}\right).$$

- Obviously, the magnitude of η tends to grow with h_t.

- The plot on p. 772 uses $n = 1$ to illustrate our points for a 3-day model.

- For example, node $(1, 1)$ of date 1 and node $(2, 3)$ of date 2 pick $\eta = 2$.

---

[a]Lyuu and Wu (R90723065) (2003).
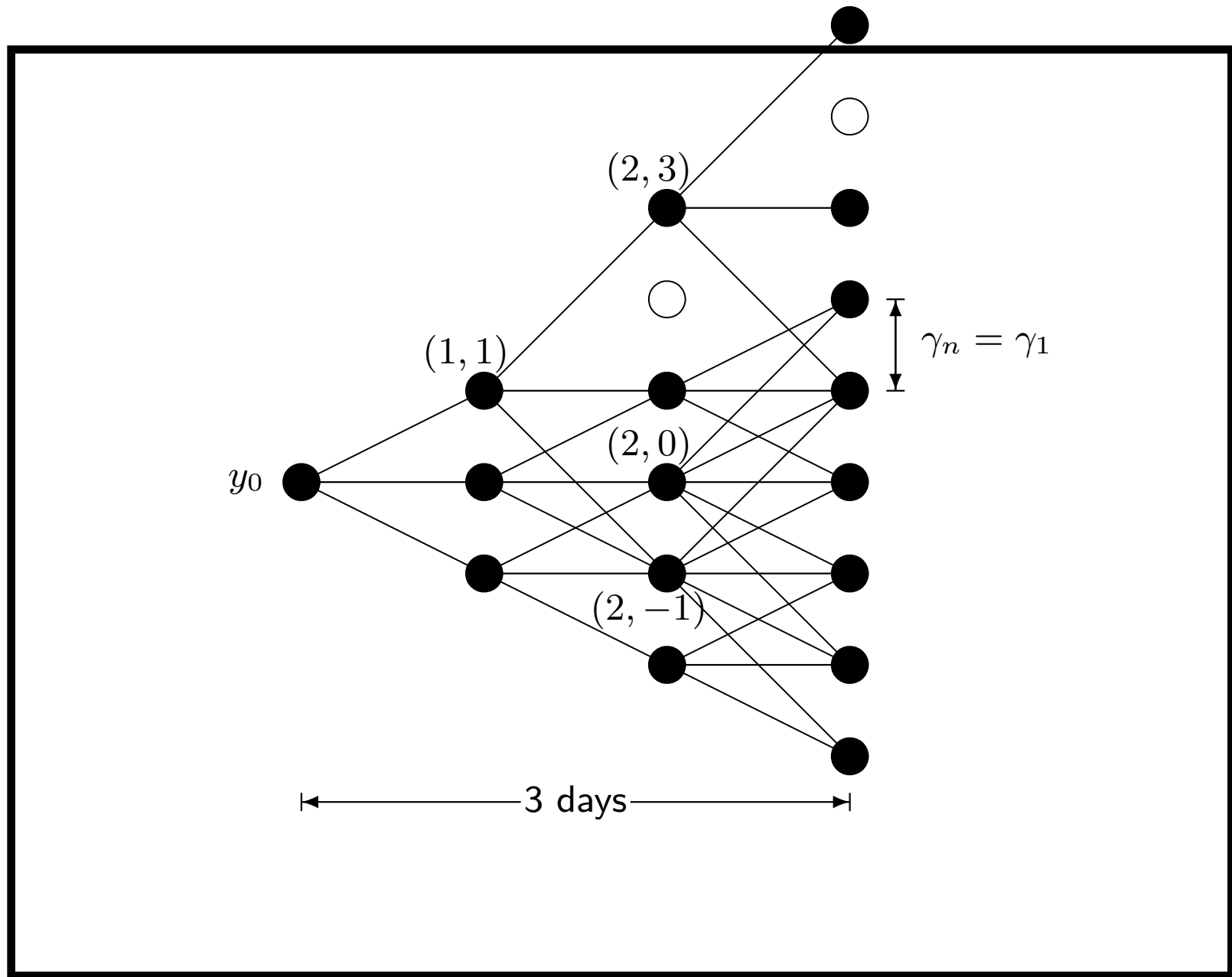
The Ritchken-Trevor Algorithm (continued)

- The topology of the tree is not a standard combining multinomial tree.

- For example, a few nodes on p. 772 such as nodes $(2, 0)$ and $(2, -1)$ have multiple jump sizes.

- The reason is the path dependence of the model.

    - Two paths can reach node $(2, 0)$ from the root node, each with a different variance for the node.

    - One of the variances results in $\eta = 1$, whereas the other results in $\eta = 2$.

The Ritchken-Trevor Algorithm (concluded)

- The number of possible values of h_t^2 at a node can be exponential.

  - Because each path brings with it a different variance h_t^2.

- To address this problem, we record only the maximum and minimum h_t^2 at each node.[a]

- Therefore, each node on the tree contains only two states $(y_t, h_{\max}^2)$ and $(y_t, h_{\min}^2)$.

- Each of $(y_t, h_{\max}^2)$ and $(y_t, h_{\min}^2)$ carries its own η and set of $2n + 1$ branching probabilities.

---

[a]Cakici and Topyan (2000). But see p. 808 for a potential problem.

Negative Aspects of the Ritchken-Trevor Algorithm[a]

- A small n may yield inaccurate option prices.

- But the tree will grow exponentially if n is large enough.
  - Specifically, $n > (1 - \beta_1)/\beta_2$ when $r = c = 0$.

- A large n has another serious problem: The tree cannot grow beyond a certain date.

- Thus the choice of n may be limited in practice.

- The RT algorithm can be modified to be free of shortened maturity and exponential complexity.[b]

---

[a]Lyuu and Wu (R90723065) (2003, 2005).
[b]It is only $O(n^2)$ if $n \leq (\sqrt{(1 - \beta_1)/\beta_2} - c)^2$!